SIFT

Robby T. Tan

Image Matching

Image Matching

Image Alignment/Registration with SIFT

Image Alignment/Registration with SIFT

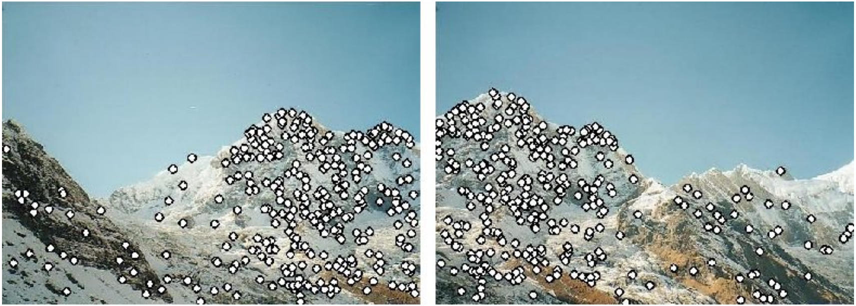


Image Alignment/Registration with SIFT

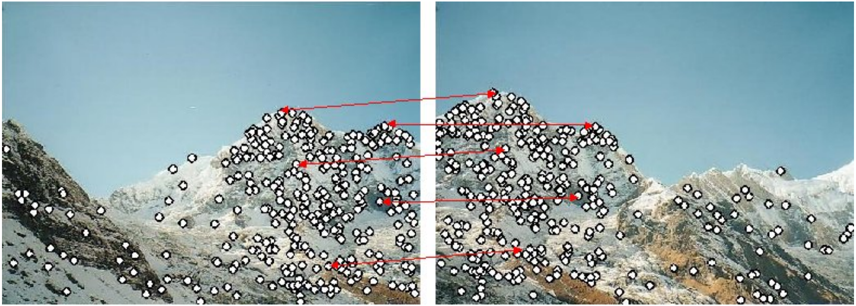


Image Alignment/Registration with SIFT

SIFT: Basic Steps

1. Detection of Scale-Space Extrema
2. Keypoint Localization
3. Orientation Assignment
4. Local Image Descriptor

1. Detection of Scale-Space Extrema

Purpose: to identify locations and scales that can be repeatly assigned under different scene conditions of the same object.

Scale Space

Scale-space problem is the problem when we see the same object with different distance.

When we change the distance, there are at least two things happening:

1. the size of the object changes
2. the details of the object change

In the scale space, the first is solved by creating **pyramid images**,
the second is by **convolving with Gaussian functions**.

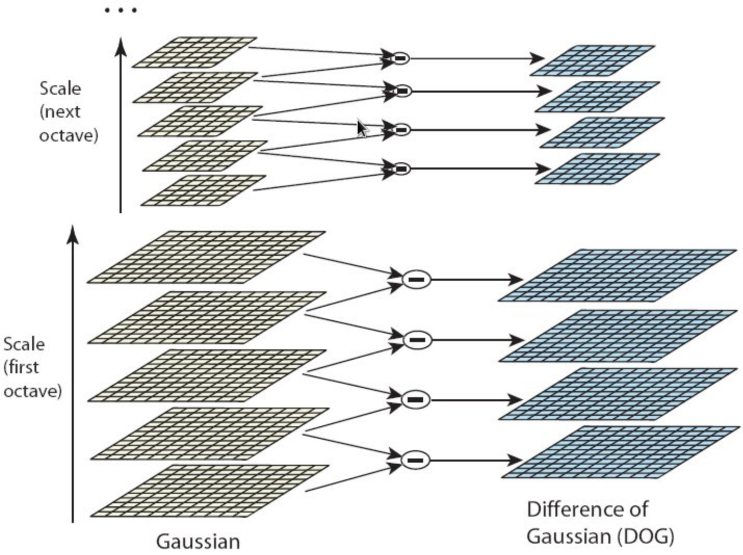
Pyramid Images (Image Scaling)

The size of the objects changes

Gaussian Convolution

The details of the objects change

Difference of Gaussian



Gaussian Convolution Over Different Scales

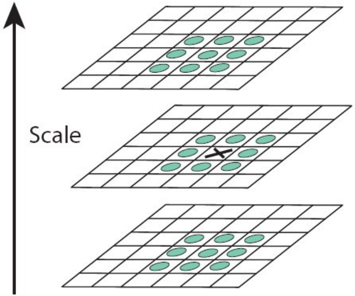
Difference of Gaussian

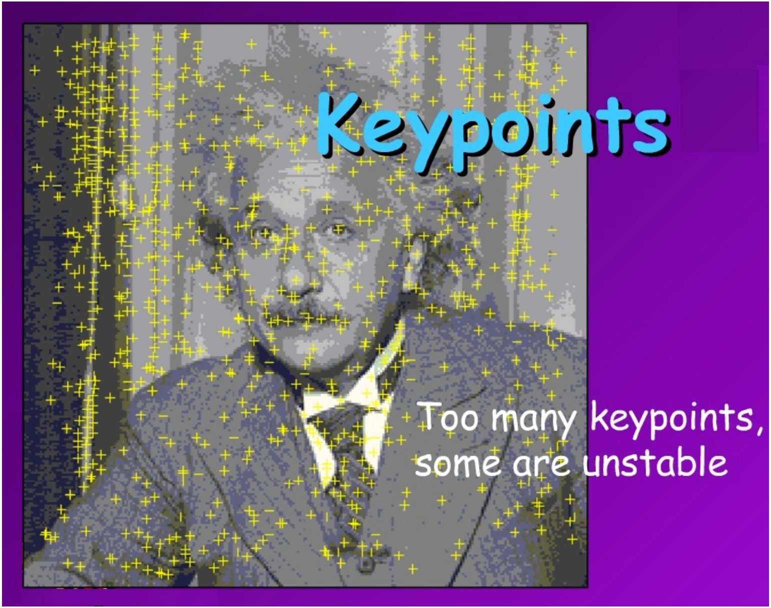
Local Extrema Detection

Compare each pixel to:

1. 8 neighbors in current image
2. 9 neighbors in scale above
3. 9 neighbors in scale below

Take pixel as a keypoint candidate if it is larger than all of them.




Keypoints
Too many keypoints,
some are unstable

2. Keypoint Localization

Two basic operations:

1. Reject points with low contrast
2. Reject points that are localized along an edge

2.a: Low Contrast Rejection

Use Taylor expansion (up to quadratic terms) of the scale-space function, so that the origin is at the candidate keypoint:

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x} \tag{1}$$

where: $\mathbf{x} = (x, y, \sigma)^T$.
The location of the extremum, $\hat{\mathbf{x}}$, is determined by taking the derivative of $D(\mathbf{x})$ w.r.t. $\mathbf{x}$:

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2}\frac{\partial D}{\partial \mathbf{x}} \tag{2}$$

if the offset $\hat{\mathbf{x}} > 0.5$: the extremum lies closer to a different sample point!

2.a: Low Contrast Rejection

Calculate D at the extremum point $\hat{\mathbf{x}}$:

$$D(\hat{\mathbf{x}}) = D + \frac{\partial D^T}{\partial \hat{\mathbf{x}}} \hat{\mathbf{x}} \tag{3}$$

If $|D(\hat{\mathbf{x}})| < 0.03$ discard the keypoint for having a low contrast.

2.b: Edge Elimination

DoG function might have strong response along edges.

We intend to remove these edges by identifying large principal curvature across the edges, yet a small one in the perpendicular direction.

2.b: Edge Elimination

1. Compute the Hessian matrix:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{4}$$

2. Compute the trace and the determinant of H:

$$\begin{aligned} \mathrm{Tr}(H) &= D_{xx} + D_{yy} = \alpha + \beta \quad (5) \\ \mathrm{Det}(H) &= D_{xx} D_{yy} - (D_{xy})^2 = \alpha\beta \quad (6) \end{aligned}$$

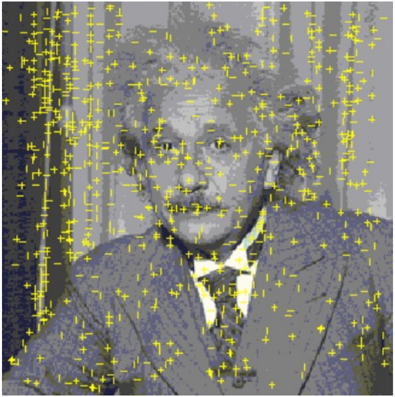
3. Compute the ratio of the trace and the determinant:

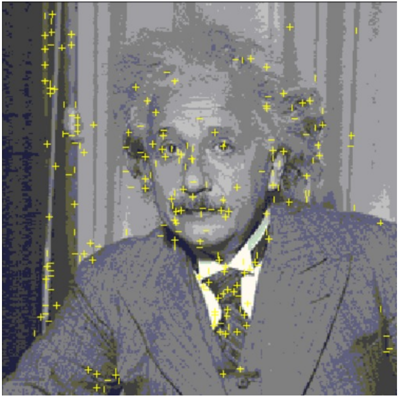
$$\frac{\mathrm{Tr}^2(H)}{\mathrm{Det}(H)} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \tag{7}$$

4. Check if the following is true when $r = 10$:

$$\frac{\mathrm{Tr}^2(H)}{\mathrm{Det}(H)} < \frac{(r+1)^2}{r} \tag{8}$$

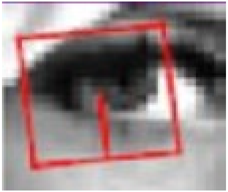
2. Keypoint Localization

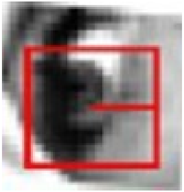




3. Orientation Assignment

Data is transformed relative to the assigned orientation, scale and location, hence providing invariance to these transformation:

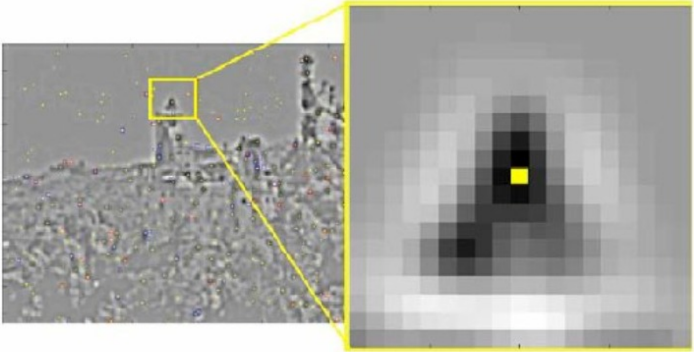




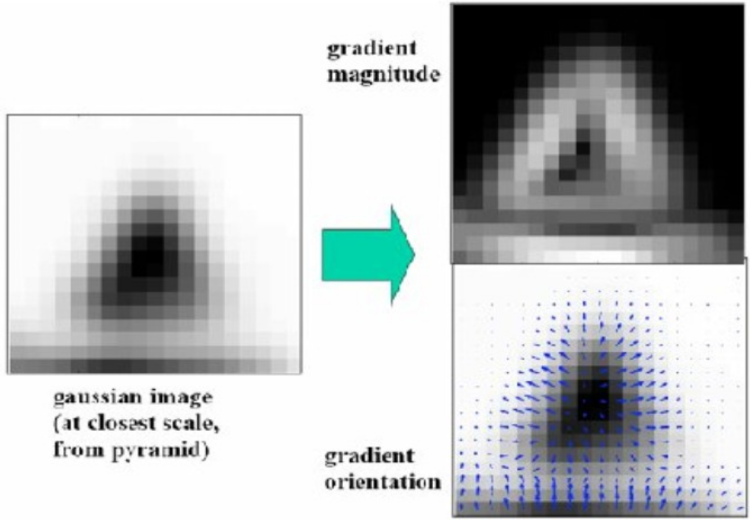
Gradient Calculation

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$

$$\theta(x,y) = tan^{-1} \frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}$$

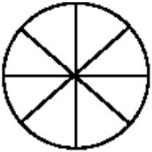


Gradient Magnitude and Orientation



Orientation Histogram

Orientation histogram with 36 bins (each bin covers 10 degrees)

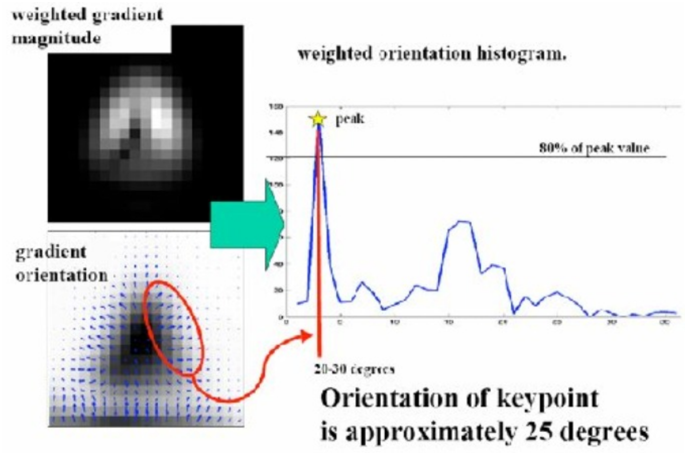


orientation histogram

36 buckets
10 degree range of angles in each bucket, i.e.
0 <=ang<10 : bucket 1
10<=ang<20 : bucket 2
20<=ang<30 : bucket 3 …

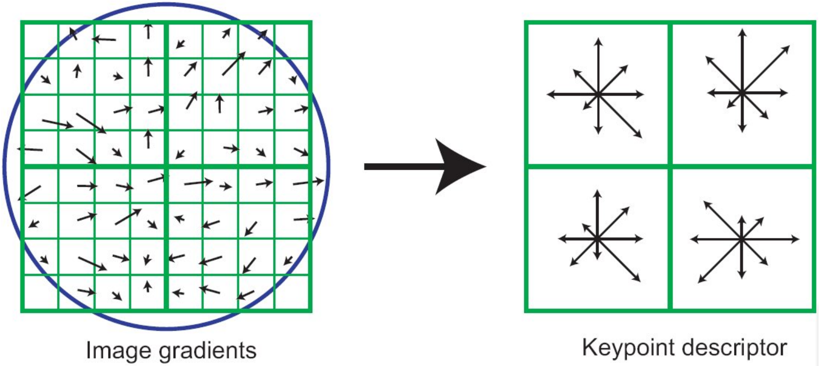
Each sample added to the histogram is weighted by its gradient magnitude and by a Gaussian-weighted circular window with a σ that is 1.5 times that of te scale of the keypoint.

Orientation Histogram



Detect the highest peak and local peaks that are within 80% of the highest peak. Use these to assign one or more orientations.

Local Image Descriptor



Local Image Descriptor